WORLD HEALTH ORGANIZATION
INTERNATIONAL AGENCY FOR RESEARCH ON CANCER

IARC Handbooks of Cancer Prevention

W O R K I N G P R O C E D U R E S

SCREENING

LYON, FRANCE
2016

Contents

HANDBOOKS OF CANCER PREVENTION – SCREENING

The Working Procedures of the *IARC Handbooks of Cancer Prevention* describe the objective and scope of the programme, the scientific principles and procedures used in developing a *Handbook*, the types of evidence considered, and the scientific criteria that guide the evaluations.

A. GENERAL PRINCIPLES AND PROCEDURES

1. Background

The global burden of cancer is high and continues to increase: the annual number of new cases was estimated at 14.1 million in 2012 and is expected to reach 22.2 million by 2030 (Ferlay *et al*., 2014). With current trends in demographics and exposure, the cancer burden has been shifting from high-resource countries to low- and medium-resource countries.

Prevention of cancer is one of the key objectives of the International Agency for Research on Cancer (IARC). Cancer prevention can be achieved by primary prevention – aimed at preventing the occurrence of cancer – or by secondary prevention – aimed at diagnosing cancer sufficiently early to reduce related mortality and suffering.

Screening and early clinical diagnosis are the principal instruments of secondary prevention of cancer and a fundamental component of any cancer control programme. Screening may enable detection of cancer sufficiently early that cure and resulting reduction in mortality and suffering from the disease are realistic possibilities given suitable treatment. Screening for some cancers, such as cervical cancer, may also detect precancerous lesions, effective treatment of which can prevent occurrence of cancer.

When screening is planned as part of a cancer control programme, only strategies proved to be effective should be proposed to the general population. Screening usually requires repeated interactions between "healthy" individuals and health-care providers, which can be inconvenient and

costly. Furthermore, screening requires an ongoing commitment between the public and health-care providers.

2. Scope

Cochrane (1972) first discussed the concepts of efficacy and effectiveness in the context of health interventions. “Efficacy” was recently defined by Porta (2008) as "the extent to which a specific intervention, procedure, regimen or service produces a beneficial result under ideal conditions; the benefit or utility to the individual or the population of the service, treatment regimen, or intervention. Ideally, the determination of efficacy is based on the results of a randomized controlled trial". In contrast, the related term "effectiveness" is defined by the same author as "a measure of the extent to which a specific intervention, procedure, regimen or service, when deployed in the field in routine circumstances, does what it is intended to do for a specific population; a measure of the extent to which a health care intervention fulfils its objectives in practice". The distinction between efficacy as measured in experimental studies and the effectiveness of a mass population intervention is a crucial one for public health decision-making. In particular, the fact that the effectiveness of a screening procedure may be different in different populations is often overlooked. A mass programme of screening must satisfy certain minimal requirements (e.g. acceptability, availability of relevant personnel, facilities for screening, and access to pertinent health services) if it is to achieve the results that have been documented in epidemiological studies.

The acceptance and use of screening services may vary from one population to another, implying that a given screening procedure is not universally effective. Even when a screening procedure is effective as a mass intervention, other outcomes, such as harm and costs and the potential for other interventions to achieve equivalent benefits, must be considered. Efficacy is a necessary but not sufficient basis for recommending screening. The efficacy of a screening procedure can be inferred if effectiveness can be proven. Screening has sometimes been implemented by a given procedure on the assumption that “earlier is better”, even when no evidence of efficacy was available. If such interventions result in a significant reduction in mortality that cannot otherwise be explained, it can be inferred that the procedure is effective.

However, uncontrolled interventions in which individuals are exposed to unknown risks and benefits should be avoided.

3. Objectives

The objectives of the Working Group are:

(1) To evaluate the strength of the evidence for the preventive efficacy of a screening procedure;
(2) To assess the effectiveness of defined screening interventions in defined populations;
(3) To assess the balance of benefit and harm in target populations;

The conclusions of the Working Group are published as a volume in the *IARC Handbooks of Cancer Prevention* series.

4. Meeting participants

Five categories of participant can be present at a *Handbook* meeting:

(1) The Working Group is responsible for the critical reviews and evaluations. The tasks of *Working Group Members* are described in detail below. Working Group Members are selected on the basis of: (i) knowledge and experience; and (ii) absence of real or apparent conflicts of interests. They have often published significant research related to the intervention being reviewed, and IARC uses literature searches to identify such experts. Experts in the general subject matter or methodology who have not published on the subject of the evaluation may also be included. Consideration is also given to demographic diversity and balance of scientific findings and views.
(2) *Invited Specialists* are experts who also have important knowledge and experience, but have a real or apparent conflict of interests. These experts are invited when necessary to assist the Working Group by contributing technical knowledge and experience during subgroup and plenary discussions. They may also review text prepared by the Working Group and contribute text on issues that do not influence the final evaluation, for example, description of the agent evaluated (for chemicals) or techniques (for screening) (see Part B, Section 2). Invited

Specialists do not serve as meeting chair or subgroup chair, and do not participate in the evaluations.

(3) *Representatives* of national and international health agencies often attend meetings because their agencies are sponsors of the programme or are interested in the subject of a meeting. Representatives do not serve as meeting chair or subgroup chair, do not draft any part of a *Handbook*, and do not participate in the evaluations.

(4) *Observers* with relevant scientific credentials may be admitted to a meeting in limited numbers. Attention will be given to achieving a balance of Observers from constituencies with differing perspectives. They are invited to observe the meeting and should not attempt to influence it. At the meeting, the meeting chair and subgroup chairs may grant Observers an opportunity to speak, generally after they have observed a discussion. Observers agree to respect the Guidelines for Observers at Meetings of the *IARC Handbooks of Cancer Prevention* (available at http://handbooks.iarc.fr).

(5) The *IARC Secretariat* consists of IARC scientists who have relevant expertise. They serve as rapporteurs and participate in all discussions. When requested by the meeting chair or subgroup chair, they may also draft text or prepare tables and analyses. They do not participate in evaluations.

Before an invitation is extended, each potential participant, including the IARC Secretariat, completes the "Declaration of Interests for IARC/WHO Experts" form to report financial interests, employment and consulting, and individual and institutional research support related to the subject of the meeting. IARC assesses these interests to determine whether there is a real or apparent conflict that warrants some limitation on participation. The declarations are updated and reviewed again at the opening of the meeting. Interests related to the subject of the meeting are disclosed to the meeting participants and in the published volume.

The names and principal affiliations of participants are available on the website of the *IARC Handbooks of Cancer Prevention* (http://handbooks.iarc.fr)

approximately two months before each meeting. It is not acceptable for Observers or third parties to contact other participants before a meeting or to lobby them at any time. Meeting participants are asked to report all such contacts to IARC.

All participants are listed, with their principal affiliations, at the beginning of each volume. Each participant who is a Working Group Member serves as an individual scientist and not as a representative of any organization, government, or industry.

5. Working procedures

A separate Working Group is responsible for developing each volume of the *Handbooks*. Approximately one year before the Working Group meeting, the agents to be reviewed are announced on the *Handbooks* website (http://handbooks.iarc.fr) and participants are selected by IARC staff in consultation with other experts. Subsequently, IARC performs literature searches of recognized sources of information on cancer prevention. Meeting participants are expected to supplement the IARC literature searches with their own searches.

The relevant articles are made available to meeting participants, who prepare preliminary drafts of the sections assigned to them. The preliminary drafts are sent to Working Group Members and Invited Specialists for peer review, and the peer-review comments are sent to the original author, who revises the draft before the meeting.

The Working Group meets at IARC for eight days to discuss and review the text and to formulate the evaluations. The objectives of the meeting are peer review, evaluation, and consensus. During the first few days, the participants meet in subgroups to review the drafts of their subgroup, develop a joint draft, and write summaries. Care is taken to ensure that each study summary is written or reviewed by someone not associated with the study being considered. During the last few days, the Working Group meets in plenary session to review the subgroup drafts and develop the evaluations. As a result,

the entire volume is the joint product of the Working Group, and there are no individually authored sections.

IARC Working Groups strive to achieve a consensus evaluation. Consensus reflects broad agreement among Working Group Members, but not necessarily unanimity. The chair may elect to poll Working Group Members to determine the diversity of scientific opinion on issues where consensus is not readily apparent.

Thus, the tasks of the Working Group are as follows:

(1) Ascertain that all appropriate data have been retrieved;
(2) Select the data relevant for evaluation on the basis of scientific merit;
(3) Prepare summaries of the data that will allow the reader to follow the reasoning of the Working Group;
(4) Evaluate separately the efficacy and the effectiveness of the screening procedure;
(5) Summarize the potential adverse consequences of screening;
(6) Prepare an overall evaluation of the screening procedure at the population level, combining all lines of evidence.

A summary of the outcome is published on the *Handbooks* programme website and as a short report in the *New England Journal of Medicine* shortly after the meeting. Subsequently, the accuracy of the final draft ("master") is verified by consulting the original literature, and the volume is edited and prepared for publication. The aim is to publish the volume within 12 months after the Working Group meeting.

6. Inclusion criteria for data for the *Handbooks*

The *Handbooks* do not necessarily summarize or even cite the entire literature on the intervention being evaluated. Only those data considered by the Working Group to be relevant to making the evaluation are included. Data judged to be inadequate or irrelevant to the evaluation may, at the discretion of the Working Group, be cited but not summarized. If a group of similar studies is not reviewed, the reasons are indicated (see Part B for details).

Meeting abstracts and other reports that do not provide sufficient detail upon which to base an assessment of their quality are generally not considered. With regard to reports of basic scientific research, epidemiological studies, clinical trials, and meta-analyses, only those that have been published or accepted for publication in the openly available scientific literature are reviewed by the Working Group. The same publication requirement applies to meta-analyses or pooled analyses commissioned by IARC in advance of a meeting (see Part B). Government agency reports that have undergone peer review and that are publicly available are considered. Exceptionally, doctoral theses and other materials that are in their final form and publicly available may be reviewed if their inclusion is considered pertinent to making a final evaluation.

B. SCIENTIFIC REVIEW AND EVALUATION

The available studies are summarized by the Working Group, with particular regard to the qualitative aspects discussed below.

Inclusion of a study does not imply acceptance of the adequacy of the study design or of the analysis and interpretation of the results. Major limitations, important aspects of a study that directly impinge on its interpretation, or reasons for not giving further consideration to an individual study are brought to the attention of the reader by the addition of square bracket comments.

Studies that are judged to be inadequate or irrelevant to the evaluation are generally omitted. They may be mentioned briefly: (i) when the information is considered to be a useful supplement to that in other reports; (ii) if they provide the only data available; or (iii) in exceptional cases, if they have been perceived as being pertinent by the scientific community but are deemed otherwise by the Working Group.

The Working Group may conduct additional analyses of the published data and use these in their assessment of the evidence. They are usually identified by square bracket comments.

The framework of a *Handbook* on screening includes the following sections.

1. Global burden and disease characteristics

Descriptive epidemiology

The purpose of this section is to document the importance of the disease in terms of the worldwide burden of the cancer described (mortality, incidence, prevalence, and survival rates), including regional differences and time trends. Expected trends in the absence of screening are a relevant component of this section.

Natural history of the disease, risk factors, treatment, and survival

In this section, the natural history of the disease of interest and the established risk factors are briefly described. Information on treatment and survival in different settings is reviewed, with a worldwide perspective.

2. Screening techniques

It is important to distinguish between screening techniques and screening procedures, i.e. between the technique itself and the way in which it is administered. The two merit separate, detailed evaluation. Each of the screening techniques to be considered is described. The ability of each test to detect cancer and to distinguish cancer from non-cancer conditions is assessed:

- Technique of screening test;
- Technical quality control;
- Screening performance;
- Host factors affecting screening performance;
- Cost of the test when implemented in mass screening programmes.

3. Availability and use of screening programmes

Information on how screening is delivered in different countries is reviewed in this section, with emphasis on the following aspects:

- Infrastructure for diagnosis and treatment: standard diagnostic procedures and treatment regimens and their availability to the target population;

- Extent of population coverage and participation rates;
- Equity, as defined by the extent to which access to the procedure (including diagnostic investigation and treatment) is ensured for all eligible individuals, irrespective of any personal characteristics;
- Informed decision and informed consent: the extent to which individual values are respected when information on potential benefit and harm is conveyed and recommendations for screening made;
- Behavioural and demographic considerations that affect participation in screening.

4. Efficacy of screening tests

In this section, evidence from efficacy studies is reviewed, and aspects of study design and analysis are critically discussed. The *Handbooks* are not intended to summarize all published studies (see Part A). The Working Group considers the following aspects:

- Relevance of the study;
- Appropriateness of the design and analysis to the question being asked;
- Adequacy and completeness of the presentation of the data;
- Degree to which chance, bias, and confounding may have affected the results.

The appropriate outcomes (mortality or incidence) of a given procedure, for example the detectable phases of the natural history of the disease, are also defined.

Aspects that are particularly important in evaluating randomized controlled trials are: the selection of participants, the nature and adequacy of the randomization procedure, evidence that randomization achieved an adequate balance between the groups, exclusion criteria used before and after randomization, compliance with the intervention in the screened group, and "contamination" of the control group with the intervention. Other considerations are the means by which the end-point was determined and validated (either by screening or by other means of detection of the disease),

the length and completeness of follow-up of the groups, and the adequacy of the analysis.

When randomized controlled trials are lacking, relevant observational studies should be considered and similar criteria used for their evaluation. In evaluating case–control and cohort studies, particular attention is paid to the definition of cases, controls, and exposure and, for cohort studies, to the length and completeness of follow-up. Potential bias, especially selection bias, is carefully examined in all observational studies.

5. Effectiveness of population-based screening

The impact of the screening procedure when implemented in defined populations is examined in this section. Indicators used to monitor effectiveness, such as positive and negative predictive values, detection rate, rates of interval cancers, and the number of tests performed, are reported. Time trends before and after implementation of screening as well as comparisons, including geographical comparisons, of the occurrence of the disease and death from the disease in populations exposed and not exposed to screening are reviewed and interpreted. In doing this, the Working Group takes into account differences in screening procedures (e.g. frequency and the age of the target population) and of participation rates.

An integral component of this section is an evaluation of the expected benefit or harm of the screening procedure to the population. Reductions in mortality from and/or incidence of invasive disease are fundamental indicators of benefit. An additional benefit is that more cases may be treated initially by less aggressive, less invasive procedures, thus improving quality of life.

The spectrum of health care is dynamic, and a screening procedure should not be viewed in isolation. Greater awareness of the disease, brought about by publicity about screening that may result in early diagnosis, could be regarded as another benefit of a screening programme. Also, in this section the possibility should be considered that there might have been a change in treatment of the cancer, which even in the absence of screening would have resulted in a substantial decrease in mortality. As far as possible, an evaluation

should be made of the extent to which improved treatment has been responsible for any changes seen in mortality from the specific disease. Estimates of rates of false-positive and false-negative findings in screened individuals and their consequences (false sense of security with false-negatives, and false alarm and consequent diagnostic and sometimes therapeutic intervention with false-positives) are an integral part of this section. The rates of short- and long-term side-effects of the screening procedure and the likelihood of unnecessary treatment are discussed.

Management procedures for lesions detected at screening are reviewed. Psychological factors, such as anxiety induced by undergoing the test procedure, are also considered. Finally, the cost–effectiveness of various modalities of test administration in various settings is considered. The discussion takes into account the costs per case detected and per death prevented.

6. Summary

In this section, the relevant data from each of the previous sections are summarized. Inadequate studies identified in the preceding text are not included.

7. Evaluation

Evaluations of the screening procedures

An evaluation of the degree of evidence of the efficacy and of the effectiveness of each screening procedure is formulated according to the following definitions.

Sufficient evidence for the efficacy and effectiveness of a cancer-preventive activity will apply when screening interventions by a defined procedure are consistently associated with a reduction in mortality from the cancer and/or a reduction in the incidence of invasive cancer, and chance and bias can be ruled out with reasonable confidence.

Limited evidence for the efficacy and effectiveness of a cancer-preventive activity will apply when screening interventions by a defined procedure are associated with a reduction in mortality from the cancer and/or a reduction in the incidence of invasive cancer, or a reduction in the incidence of clinically advanced cancer, but bias or confounding cannot be ruled out with reasonable confidence as alternative explanations for these associations.

Inadequate evidence for the efficacy and effectiveness of a cancer-preventive activity will apply when data are lacking, or when the available information is insufficient or too heterogeneous to allow an evaluation.

Sufficient evidence that the screening procedure is not efficacious in cancer prevention will apply when any of the following cases hold:

- The procedure does not result in earlier diagnosis than with standard methods already in use;
- The survival of cases detected at screening is no better than that of cases diagnosed routinely;
- The screening interventions are consistently associated with no reduction in mortality from or incidence of invasive cancer, and bias can be ruled out with reasonable confidence.

In the case of *limited* or *inadequate* evidence, the Working Group should highlight those aspects of the procedure for which information is lacking, and which led to the uncertainty in evaluation. This will provide indications of research priorities.

Overall evaluation

The body of evidence for each screening procedure is considered as a whole, and summary statements are made about the cancer-preventive effects of the screening intervention and other beneficial or adverse effects, as appropriate. The overall evaluation is usually in the form of a narrative. The data on the effectiveness of the screening intervention are summarized, including the factors that determine its success and failure under routine conditions. Finally, the balance between expected benefit and harm is described.

References

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D.D. & Bray, F. (2014) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* (in press)

Cochrane, A.L. (1972) *Effectiveness and Efficiency: Random Reflections on Health Services*. Oxford, Nuffield Provincial Hospitals Trust.

Porta, M., ed (2008) *A Dictionary of Epidemiology*, 5th edition. Oxford, Oxford University Press.